图书情报工作
第 63 卷 第 1 期 2019 年 1 月

图书情报领域人工智能的研究热点及发展趋势研究

王晰巍[1,2] 贾若男[1] 王铎[1] 郭宇[1]
[1] 吉林大学管理学院 长春 130022 [2] 吉林大学大数据管理研究中心 长春 130022

摘要：［目的/意义］搜集和分析国内外图书情报领域人工智能研究的相关文献，以期了解人工智能在图书情报领域目前的研究进展及未来的研究趋势，为学术界和产业界的研究提供借鉴和启示。［方法/过程］运用文献分析和知识图谱可视化方法，对国内外图书情报领域人工智能研究的起源和发展、国内外研究热点及未来研究趋势进行对比和可视化分析。［结果/结论］从研究发展趋势看，国内外相关研究成果的数量呈逐年递增态势。从研究热点看，国外研究主要围绕人工智能在信息检索、社交媒体内容分类及情感分析、知识问答方面的应用研究；国内研究主要围绕人工智能在网络舆情分析、图书馆智能搜索和推荐服务、社交媒体网络数据分析方面的研究。从研究趋势看，未来研究应围绕人工智能技术应用在视觉搜索和语义搜索、智慧图书馆服务、社交网络数据挖掘、人工智能对隐私与安全的影响 4 个主要方向来开展研究。

关键词：图书情报 人工智能 研究热点 发展趋势
分类号：TP18
DOI：10.13266/j.issn.0252－3116.2019.01.011

1 引言

人工智能(artificial intelligence，AI)是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学[1]。麦肯锡《2017 全球人工智能报告》中指出人工智能将造成下一波数字化颠覆，像谷歌和百度这样的数字巨头公司正在主导着人工智能投资的快速增长，人工智能正在成为世界领先的科技公司之间的专利和知识产权的重要竞争要素[2]。同样，从猎豹智库发布的《2018 中国人工智能报告》来看，国内人工智能在应用层面创新加速的条件已经趋于成熟，再加上相关政策的引导和各方资本的支持，我国人工智能在各垂直领域的应用已经全面绽放[3]。

国内外与人工智能相关的研究涉及多个学科领域，其中相关研究成果最多的是在计算机软件及计算机应用领域，其次是自动化技术领域。在大数据和人工智能快速发展的时代背景下，大数据、机器学习以及深度学习等人工智能的算法和技术，在图书情报领域的研究中发挥越来越重要的作用。同时，人工智能强大的分析能力和数据处理能力也为图书情报领域的研究带来了新的机遇和挑战。研究分析人工智能在图书情报领域的应用热点及未来发展趋势，对推动人工智能在图书情报领域的应用具有重要的理论和现实意义，也成为图书情报领域新的研究课题。

本文在研究中试图研究以下 3 个方面的问题：①人工智能在国内外图书情报领域研究的起源与发展如何，各自有什么异同；②近几年人工智能在国内外图书情报领域的研究热点；③人工智能在图书情报领域未来的研究趋势。本文对人工智能在国内外图书情报领域应用研究的相关文献进行梳理，并通过 CiteSpace 软件绘制知识图谱，对人工智能在国内外图书情报领域中的研究热点和发展趋势进行分析，以期为国内外图书情报领域学者对人工智能在图书情报领域的深入理论研究和实践探索提供参考和借鉴。

2 样本选择

为确定样本选择中的关键词，依据《中国人工智能报告 2018》[4]，确定人工智能领域的关键词为人工智能、机器学习与深度学习，同时该报告显示与 AI 相关的具体技术还有自然语言处理、图像识别、人脸识别、

作者简介：王晰巍(ORCID：0000－0002－5850－0126)，副院长，大数据管理研究中心主任，教授，博士生导师，E-mail：wxw_mail@163.com；贾若男(ORCID：0000－0002－4262－7982)，硕士研究生；王铎(ORCID：0000－0002－5060－7893)，博士研究生；郭宇(ORCID：0000－0003－1782－7752)，副教授，博士。
收稿日期：2018－10－16 修回日期：2018－11－16 本文起止页码：70－80 本文责任编辑：王传清

计算机视觉、虚拟助手,视觉搜索、文本分析、语义搜索、预测分析、智能系统以及语义网络等。考虑检索结果的全面性,在制定检索式时将上述关键词一同列入其中。

在国外文献样本的选择上,选择文献质量和收录范围具有一定优势的 Web of Science 数据库(Web of Science 包含 SSCI 和 SCI 两个数据库),并以 TS(主题) =("artificial intelligence") or TS =("machine learning") or TS =("deep learning") or TS =("natural language processing") or TS =("image recognition") or TS =("face recognition") or TS =("computer vision") or TS =("virtual assistant") or TS =("visual search") or TS =("text analysis") or TS =("semantic search") or TS =("predictive analysis") or TS =("intelligent system") or TS =("semantic web")为检索式,检索时间段截至 2017 年 12 月 31 日,共检索到相关文献 175 376 篇,涉及计算机科学、工程学、自动化、数理学和医学等学科领域,将研究领域限定在 Information Science & Library Science,共得到相关研究文献 2 399 篇。最后依据期刊的影响因子及论文的研究内容,人工筛选后确定用于热点分析的关键文献为 35 篇。

在国内文献的选择上,本文选择集期刊杂志、学位论文、会议论文等内容资源丰富且优质的中国知网,并对相关文献进行检索,检索式为:SU(主题) =("人工智能")or SU =("机器学习")or SU =("深度学习")or SU =("自然语言处理")or SU =("图像识别")or SU =("人脸识别")or SU =("计算机视觉")or SU =("虚拟助手")or SU =("视觉搜索")or SU =("文本分析")or SU =("语义搜索")or SU =("预测分析")or SU =("智能系统")or SU =("语义网络"),检索时间段截止 2017 年 12 月 31 日,共检索到文献 109 630 篇,涉及自动化技术、计算机软件及计算机应用、工业经济等学科领域,将研究领域限定在图书情报领域后,得到相关研究文献 911 篇。最后依据图书情报领域期刊的影响因子及论文的研究内容,人工筛选后确定用于热点分析的关键文献为 32 篇。

3 图书情报领域人工智能研究的起源与发展

3.1 国外图书情报领域人工智能研究的起源与发展

通过对国外图书情报领域人工智能研究相关文献进行归纳总结,以时间(年)为横坐标,文献数量(篇)为纵坐标,绘制国外图书情报领域人工智能研究的发展脉络图,如图 1 所示:

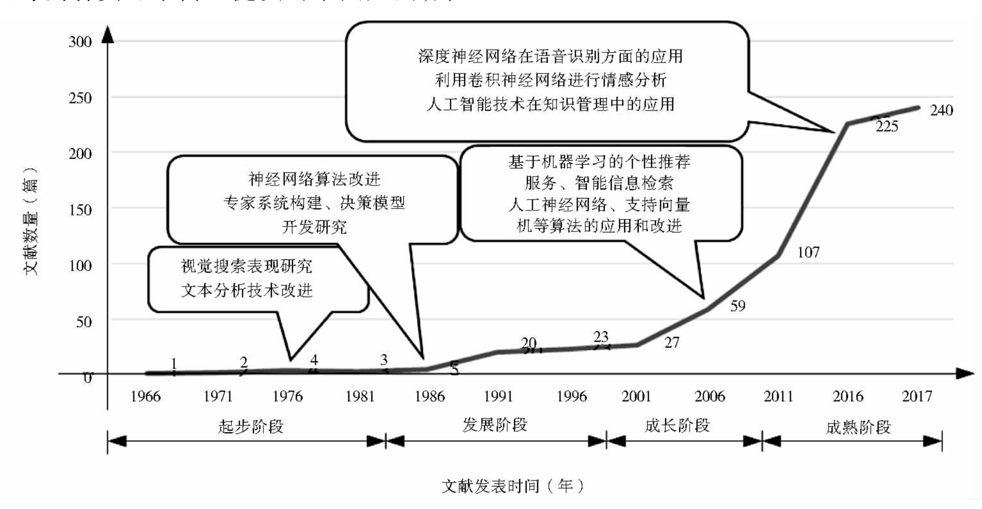


图 1 国外图书情报领域人工智能研究文献统计

3.1.1 起步阶段 国外图书情报领域的人工智能研究开始于 1966 年。从发展阶段看,1966-1983 年间,相关研究处于萌芽阶段。当时受限于信息科技的发展,每年的研究文献数量只有个位数,并且上升缓慢。这一阶段的研究主要集中在视觉搜索方面。视觉搜索的研究主要关注搜索表现,且当时的研究主要通过实验法进行。国外有学者通过设计实验,评估了目标词、上下文项和指令的变化对用户视觉搜索表现的影响[5];通过提出视觉搜索任务的熟练表现模型,用实验方法对模型进行改进[6]。此外,也有运用文本分析进行的相关研究,但数量相对很少,这些研究多聚焦于文本分析技术,还有学者提出一种名为歧视值分析的新

技术,并通过实验结果验证该技术的有效性[7]。

3.1.2 发展阶段 1983－2000 年间,国外图书情报领域在人工智能发面的研究开始进入发展初期,文献数量开始出现上升,但增幅并不明显。这一阶段的研究主要集中在神经网络和专家系统等图书情报领域应用方面的研究。神经网络的相关研究主要在算法的改进和创新融合方面,有学者构建了动态神经网络系统及其连接权的学习算法,并提出了动态 BP 参数的在线调节方法[8];也有学者以乘积 Sigmoidal 函数作为激发函数,提出了非线性系统用神经网络逼近的构造性描述,并且还估计了节点数目的上界值[9],为后续研究提供了参考。这一时期,国外学者对专家系统的研究主要围绕系统的开发,如仿照专家决策过程的 HOMBRE 决策模型,以及斯坦福研究所开发的 PROSPECTOR 系统等[10]。

3.1.3 成长阶段 2000－2010 年是图书情报领域人工智能研究的成长阶段,文献数量较前一阶段有显著增幅。随着人工智能技术的不断发展,开始出现强人工智能和弱人工智能两个不同层面的研究。强人工智能是指未来人们将被具有人类智能的机器所包围,这种机器可以完全像人类一样工作;弱人工智能是指人们可以通过各种技术将与人类活动类似的功能等添加到机器中,使机器成为更加智能的工具。这个阶段的研究主要集中在机器学习、人工神经网络、遗传算法等方面。这一时期机器学习的研究主要围绕着个性化推荐服务、智能信息检索和自动文本分类研究 3 个方面[11],主要关注人工神经网络、基于遗传算法的神经网络、支持向量机等算法和技术的优化等在图书情报领域应用[12]。

3.1.4 成熟阶段 2011 年至今,随着 Web.2.0、大数据以及互联网＋的发展,信息技术在各领域的应用更为广泛,国外图书情报领域人工智能研究也步入了成熟阶段。这个时期主要的研究方向有深度学习、情感分析、知识问答等。深度学习在语音识别、图像识别、自然语言处理等方面不断取得突破,在句法成分分析、情感分析等问题的研究上,深度学习也发挥出不错的效果[13]。国外有学者利用深度神经网络在语音识别方面不断取得突破,使得 LSTM-RNN 方法在语音识别中占据了统治地位[14];还有学者在卷积神经网络中引入动态 k-max 池化技术来处理自然语言变长的特性,在情感分类任务上取得了不错的效果[15]。人工智能技术的发展和推广应用对知识问答产生了一定的影响和支持,国外研究主要关注人工智能技术在知识问答中的应用,寻找知识问答社区中的潜在专家有利于解决用户参与率低、接收答案等待时间过久以及答案质量低等问题。有学者构建了预测知识问答社区专家排名的模型,并通过实验验证了该模型的有效性[16]。同时,协作编辑在知识问答网站中发挥着重要的作用,能够显著提高问题和答案的质量,采用 LDA 和 SVM 等机器学习技术,可以预测问题和答案是否需要在发布之后进行协作编辑,选择 Stack Overflow 平台的数百万个问答实例作为样本可以进行验证性实验[17]。

3.2 国内图书情报领域人工智能研究的起源与发展

通过对国内图书情报领域人工智能研究相关文献的归纳和总结,绘制国内图书情报领域人工智能研究发展脉络图,如图 2 所示:

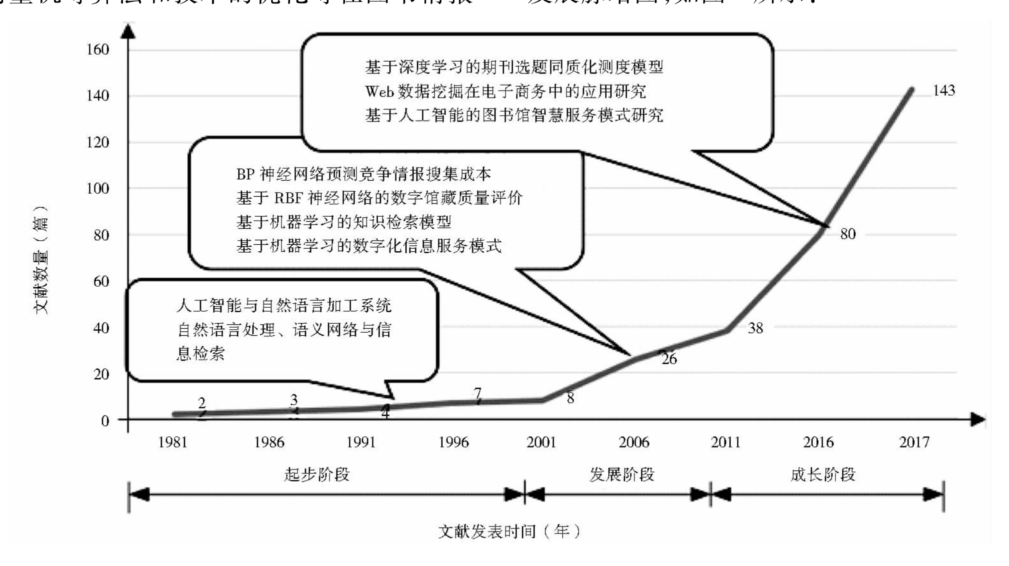


图 2 国内图书情报领域人工智能研究文献统计

3.2.1 起步阶段 国内图书情报领域的人工智能研究开始于1981年,由于科学技术的制约和研究经验的缺少,直到2000年左右一直处于低速发展状态。这一时期的研究主要集中在自然语言处理和人工智能在信息检索中的应用等,多数的研究从理论层面展开,也有少数基于实验方法的研究。有学者从理论层面论述了自然语言加工的重要性和复杂性,并提倡研制融入人工智能技术的自然语言的信息加工系统[18];讨论自然语言处理应用于全文检索的若干问题,对标引、分析检索的处理手段进行总结[19];分析基于语义网络的中文搜索引擎概念模型,并通过实验分析方法验证了该模型的可行性和有效性[20]。

3.2.2 发展阶段 2001－2010年国内图书情报领域的人工智能研究得到进一步发展,这一时期文献数量开始出现增长,主要的研究方向为神经网络和机器学习在实践领域的应用研究。一些学者尝试引入BP神经网络对企业竞争情报搜集成本进行预测,并验证了线性回归分析法与BP神经网络相比,BP神经网络的预测精度更高[21];后来又有学者提出径向基函数神经网络RBF(Radial-Basis Function)的自学习和自适应等特征,构建了基于RBF神经网络的数字馆藏质量评价模型,并通过实证研究验证了模型的正确性和可行性[22]。图书情报领域关于机器学习的研究主要集中在智能信息检索和自动文本分类等方面,有学者在综合人工智能和信息管理等理论与技术的基础上,探讨了获取检索知识的机器学习方法,并提出了基于机器学习的知识检索模型[23];也可以通过对个性化数字信息服务特点和模式的分析,构建基于机器学习的主动式自适应服务模式[24]。

3.2.3 成长阶段 自2011年起,随着新一代信息技术的快速发展,人工智能在图书情报领域的应用有了强有力的技术支持和制度保障,图书情报领域对人工智能的研究进入快速成长阶段,这个时期国内图书情报领域对人工智能的研究开始大幅度增多,研究方向也开始进入纵深化发展。图书情报领域的专家学者在延续机器学习的基础上,向深度学习、数据挖掘以及智慧服务等方面延伸,人工智能的研究开始越来越多地关注在产业实践层面的应用。深度学习的出现,使得人工智能领域中传统机器学习方法的困境得到了改善,拓宽了人工智能在图书情报领域的研究范围。有学者提出了基于深度学习的学术期刊选题同质化测度模型,分析了当前学术期刊发展中的同质化问题,并通过实证研究方法验证了基于深度学习所构建模型的有效性和可行性[25]。同时将Web数据挖掘技术应用于电子商务中,使电子商务中的智能服务具有更强的智能化和个性化[26]。通过对比传统数据挖掘和大数据挖掘,将云计算与数据挖掘服务相结合构建大数据挖掘体系架构,为大数据挖掘的认知和应用提供了参考[27]。智慧服务方面的研究主要体现在人工智能技术和理念在智慧图书馆中的应用,基于对大数据时代读者个性化阅读风险与需求的研究,构建基于大数据的图书馆个性化智慧服务体系,为智慧图书馆建设提供借鉴和参考[28]。

3.3 国内外图书情报领域人工智能研究起源与发展对比分析

从演进发展的时间段来看,国内图书情报领域对人工智能的研究滞后国外近15年,但进入2000年之后,国内图书情报领域的人工智能研究进入快速发展和成长阶段;国内外图书情报领域的人工智能研究在发展演进过程中,在神经网络、机器学习和深度学习等领域的研究内容具有一定的相似性。比较而言,近几年,国外的研究主要围绕着机器学习、深度学习算法的优化和应用,如深度神经网络、卷积神经网络在图像识别、情感分析、自然语言处理等方面的应用;国内在深度学习、卷积神经网络方面的研究近几年才开始起步,更加偏向于对已有技术或算法的阐述和比较,以及图情相关领域中的应用分析,如相关文献的计量统计与可视化分析、人工智能在智慧图书馆、知识社区中的智能问答等智慧服务中的应用等。近五年,国内图书情报领域的人工智能研究发展迅速,尤其是深度学习在情报领域的发展。

3.4 国内外图书情报领域人工智能研究人员分析

利用知识图谱对国内外图书情报领域人工智能研究的相关专家学者进行分析,数据分析选取的时间段为1966年1月1日至2017年12月31日,数据分析结果如表1和图3、图4所示:

表1 国外及国内图书情报领域人工智能研究人员发文情况

序号	国外高产作者		国内高产作者	
	发文篇数	作者	发文篇数	作者
1	12	H. Xu	11	朱庆华
2	12	J. C. Denny	7	赵宇翔
3	5	T. Lingen	6	张兴旺
4	5	C. H. Marcondes	4	焦玉英
5	5	Q. Li	4	许鑫
6	5	N. Elhadad	4	赵蓉英
7	4	M. A. Goncalves	4	韩玺
8	4	M. Guerrini	4	张璐
9	4	D. W. Bates	3	李国秋
10	4	M. Jiang	3	曾子明

国外图书情报领域在人工智能方面的研究，从论文产出数量来看排在第一位的是 H. Xu 和 J. C. Denny，他们主要关注运用自然语言处理和机器学习等技术对大数据进行分析与可视化，除此之外 J. C. Denny 还关注社交网站用户数据的挖掘和分析。T. Lingen 主要关注人工智能算法在网络信息表示、检索中的应用；C. H. Marcondes 主要关注基于语义 Web 技术以及链接开放数据技术的知识组织；Q. Li 主要关注基于主题级最大熵模型在社会情感分类中的应用以及自然语言处理技术在信息提取中的应用。

国内图书情报领域人工智能研究排名前 3 的作者依次为朱庆华、赵宇翔和张兴旺。朱庆华主要关注人工智能技术在社交网络用户行为分析中应用，以及移动视觉搜索在图书馆的应用；赵宇翔除关注移动视觉搜索在图书馆应用外，还关注国内外移动视觉搜索工具的比较分析，以及基于语音互动的知识问答社区研究；张兴旺主要关注智能信息推荐、移动视觉搜索以及云计算在图书馆的应用研究等。

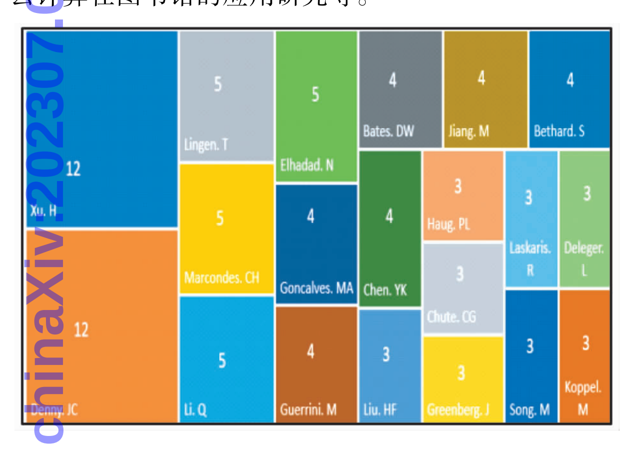


图 3　国外图书情报领域研究人员发文情况

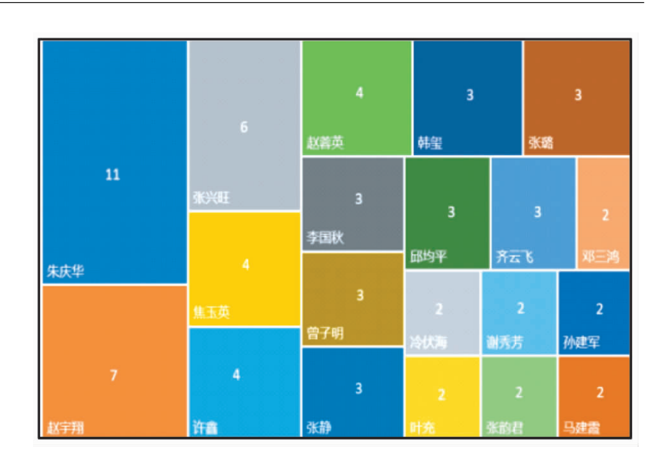


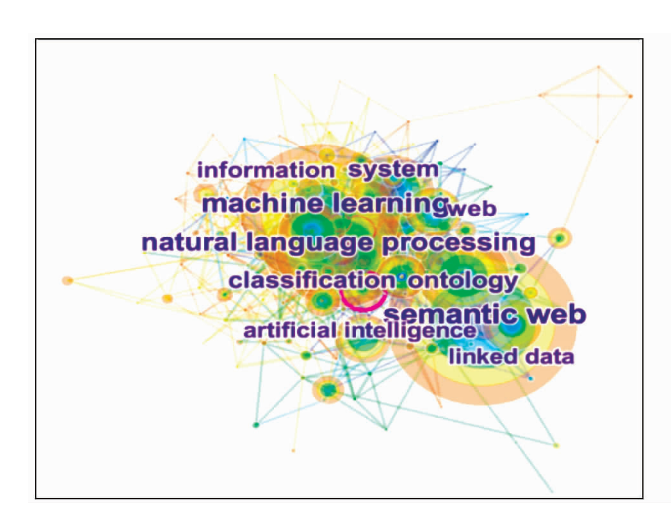
图 4　国内图书情报领域研究人员发文情况

4　国内外图书情报领域人工智能研究热点分析

4.1　国内外研究热点知识图谱分析

为了对国内外图书情报领域人工智能研究的热点进行分析，本研究采用 CiteSpace 软件对近 5 年国内外图书情报领域人工智能研究热点绘制关键词知识图谱，将 Citespace 中的 Time Slicing 设置为 2013 年 – 2017 年，Node Types 设置为 Keyword，Selection Criteria 设置为 T50，得到国内外研究热点的知识图谱如图 5 所示，并对国内外图书情报领域人工智能研究热点的高频关键词进行统计分析，见表 2。

图 5 中，圆环代表关键词节点，节点的大小反映关键词的影响力，节点越大表示出现的频次越高，影响力也越大。节点之间的连线代表两个关键词共同出现的次数，线条越粗，共同出现的频次越高。图中关键词字体的大小及圆环大小客观上反映了不同时期的研究热度，表 2 的中心度则反映了关键词节点的重要性程度。



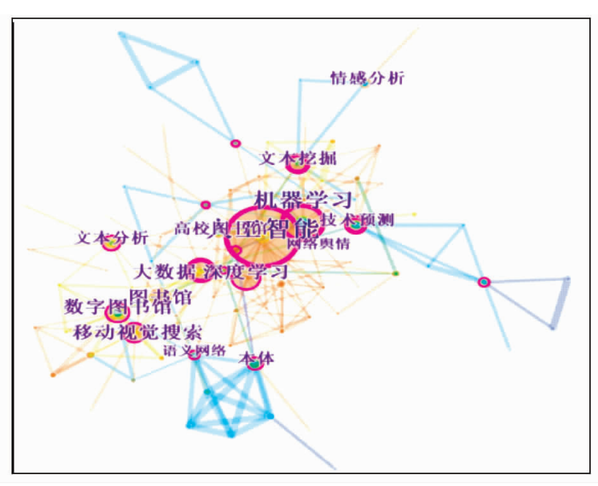


图 5　国外(左)及国内(右)图书情报领域人工智能研究热点知识图谱

观察与分析表 2 和图 5 可知,国外图书情报领域人工智能研究相关文献按关键词被引频次排序依次为 semantic web、machine learning、natural language processing、classification、ontology、system、information、artificial intelligence、linked data、information retrieval 等,这反映了近 5 年国外相关研究的热点。按照中心度排名前 3 位的是 machine learning(0.09)、classification(0.09)和 ontology(0.09),说明这 3 个节点具有较高的重要性。国内图书情报领域人工智能研究相关文献按关键词被引频次排序依次为人工智能、机器学习、深度学习、大数据、移动视觉搜索、文本挖掘、文本分析、技术预测、情感分析和智慧图书馆,这反映了近 5 年国内相关研究的热点。按照中心度来看,排名前 3 位的有人工智能(0.34)、机器学习(0.33)和技术预测(0.26),说明这 3 个节点重要性程度较高。

通过对比分析可以发现,国内外图书情报领域人工智能研究相关文献的关键词差异较大,研究热点也有所不同。除了在机器学习和神经网络等方面的研究存在相同之外,国外研究更加偏向算法、数据挖掘等方面,如语义网、自然语言处理等;国内的研究则更加偏向于人工智能技术在实践中,比如在信息搜索、智慧图书馆建设中的应用。这一结果表明,如何将已有的人工智能技术和算法运用到实践中去,如何将人工智能与图书情报领域的实践研究相结合可能会成为未来研究发展的方向。同时对算法和技术的改进和优化也需要我们借鉴国外相关研究。

表 2　国外及国内图书情报领域人工智能研究高频主题词

序号	国外高频主题词			国内高频主题词		
	被引频次	中心度	关键词	被引频次	中心度	关键词
1	141	0.04	semantic web	45	0.34	人工智能
2	116	0.09	machine learning	28	0.33	机器学习
3	99	0.03	natural language processing	18	0.18	深度学习
4	70	0.09	classification	16	0.23	大数据
5	70	0.09	ontology	15	0.15	移动视觉搜索
6	65	0.08	system	13	0.22	文本挖掘
7	52	0.06	information	12	0.10	文本分析
8	45	0.02	artificial intelligence	12	0.26	技术预测
9	45	0.00	linked data	10	0.04	情感分析
10	44	0.02	information retrieval	5	0.02	智慧图书馆

4.2 国外图书情报领域人工智能研究热点

4.2.1 人工智能在信息检索中的应用研究　此研究热点主要是围绕神经网络、语义关联等人工智能技术在信息检索过程和改进检索表现等方面进行的研究。有学者以“Apple's Growth”为例,通过神经网络分析检索信息与关键词之间的相似性,研究了人工智能信息检索技术在图书馆资源检索中的应用[29];在智能数据分析和挖掘的背景下,运用图形识别技术,将书目信息表示为网络和节点,提出一个基于图形的交互式书目信息检索系统,改善了书目信息检索的检索表现[30];通过将时间特征纳入信息检索技术考虑范围,可构建基于关键字的链接数据信息检索框架,并通过实验可证明该框架的有效性[31]。

4.2.2 人工智能在社交媒体用户内容分类以及情感分析中的应用研究　通过使用监督式机器学习方法可以对用户在社交媒体上发布的内容进行分析,并使用来自社交媒体的 4 000 条推文作为训练集进行模型和算法训练,研究结果表明线性支持向量机和朴素贝叶斯模型优于逻辑回归分类器[32];还有学者用社交媒体数据分析目前管理人员面临的重大挑战,提出了一种基于智能技术的社交媒体分析模型,并提出多种社交媒体数据分析的方法,如通过机器学习来进行情感分析、利用图像和语音识别技术来分析社交媒体视频数据[33]。随着社交媒体的快速发展,用户在使用社交网络过程中会生产内容,这些内容不仅包含事件信息,还表达了用户的情感,因此有学者基于人工智能中的深度学习方法,设计了分析预测社交媒体用户情感轨迹的系统,并通过深度学习提高该系统情绪分析和预测的准确性[34]。

4.2.3 人工智能在知识问答中的应用研究　通过将人工智能技术应用于知识问答社区,可以提高知识问答社区用户参与率,缩短答案接收时间和提高答案质量。知识问答社区中权威用户的检测是目前面临的一个重要挑战,这个问题的解决有利于提高知识问答社群中用户参与率低,接收答案等待时间过久以及答案质量低等相关问题。因而,有学者基于信念函数提出

了一种衡量知识问答社区用户信誉度的模型,并通过实验验证了该模型可以用于改进目前对知识问答社区专家的识别[35]。此外,还有学者通过逐点学习和成对学习法构建预测知识问答社区专家排名的模型,并通过实验验证了该模型的有效性[21]。同时,协作编辑能够显著提高问题和答案的质量,通过采用三层贝叶斯概率模型(Latent Dirichlet Allocation,LDA)和支持向量机(Support Vector Machine,SVM)等机器学习技术,可以预测问题和答案是否需要在发布之后进行协作编辑,选择 Stack Overflow 平台的数百万个问答实例作为样本验证了该模型的有效性[22]。

4.3 国内图书情报领域人工智能研究热点

4.3.1 人工智能技术在网络舆情分析中的应用研究

人工智能技术在网络舆情分析中的应用研究主要围绕着舆情传播规律、舆情监测与预警以及舆情情感分析等方面进行。有学者通过语义关联分析和可视化分析软件,对学术网络舆情信息的传播规律进行研究,为社交网络环境下学术网络舆情的传播提供新的研究视角和分析方法[36];也有学者提出基于灰色关联计算的网络舆情监测模型来实现网络舆情热点事件的预测和发现,并通过仿真实验验证和分析模型的效能[37];还有学者从认知情感角度建立感情规则,并利用深度学习中的卷积神经网络模型进行建模和分析,对网络舆情事件进行情感分类,实验结果证明卷积神经网络模型在舆情情感分析中的效果优于传统模型[38]。

4.3.2 人工智能对图书馆智能搜索和推荐服务的影响研究 以文化遗产为例,有学者对视觉搜索在数字图书馆领域的应用进行了描述和分析,主要包括基本原理、体系架构与服务模块等方面[39];还有学者提出具有语义发现功能的移动视觉搜索方法,对实现智慧图书馆视觉资源内容和语义信息的全搜索具有重要意义[40];在对人工智能与信息推荐融合发展研究历程进行综述的基础上,信息智能获取可以帮助实现从“智能生产→知识智能认知→知识智能体验→知识智能推送”这一图书馆信息智能推荐业务链的演进及构建[41]。考虑到图书馆用户兴趣点受时间和频率共同影响,有学者提出了一种基于本体相似度和时间衰减的动态个性化推荐算法,并通过实验检验了该算法的普适性和有效性[42]。

4.3.3 人工智能在社交媒体网络数据分析中的应用研究 从社交网络用户在线评论的属性特征出发,采用 Python 语言实现了评论文本情感极性量化分析,与传统的机器学习算法相较而言,该算法的准确性有显著提升[43]。在传统近似字符串匹配技术的基础上,使用属性值结合社交媒体中的链接和文本数据的新方法,利用 Facebook 和 Twitter 数据作为实验数据集,可以证明该方法识别跨社交媒体相同用户特征的准确性[44]。随着社交网络产生的大量复杂数据信息覆盖了用户个性化需求,为了借助社交网络数据分析实现对用户偏好的准确预测,有学者提出一种基于上下文感知的移动社交网络用户偏好识别和预测模型,并通过实验验证了该方法的有效性[45]。

5 未来研究趋势

利用 CiteSpace 绘制近 5 年(2013 - 2017 年)此领域的知识图谱,如图 6 和图 7 所示。从研究成果的发展趋势来看,国外图书情报领域在人工智能方面的相关研究呈现从“机器学习、语义网、自然语言处理→数字图书馆、情感分析、引文分析→可视化、知识管理、社交媒体→大数据、数据挖掘、智慧图书馆→深度学习”的演进发展态势;国内相关研究呈现从“机器学习、网络舆情→数字图书馆、文本挖掘、神经网络→人工智能、大数据、语义搜索→深度学习、移动视觉搜索→智慧图书馆、智能图书馆”的演进发展态势。总体来看,图书情报领域人工智能的未来研究趋势分析如下:

5.1 人工智能技术应用于视觉搜索和语义检索研究

随着图像和视频应用的爆炸式增长,视觉资源成为当下重要的可视信息单元,人工智能技术应用于视觉资源的检索将成为情报学领域研究的创新价值增长点,视觉搜索成为信息检索研究中新的发展趋势。此外,语义搜索在网络信息检索领域也是一个新颖且具有良好前景的研究方向。未来的研究可以从以下两方面展开:一方面,视觉搜索和语义搜索框架或模型的构建,可以考虑融合多方面的特征,如视觉资源的底层视觉特征、高层语义特征以及上下文文本信息特征等,配合机器学习和深度学习技术可以来构建搜索模型[46],以提高搜索的效率和搜索结果的准确性。另一方面,由于视觉搜索和语义搜索性能的提高,以海量网络数据作为驱动,利用机器学习和深度学习等技术,可以对已有视觉搜索模型的检索准确度进行改良和优化[47]。利用计算机视觉技术检测文本或图像数据并生成自动标签,可改善检索到信息的效率,并提高检索模型的准确性[48]。

5.2 人工智能在智慧图书馆服务中的应用研究

智慧图书馆是当下图书馆发展的新趋势[49],人工智能技术在智慧图书馆的应用研究也是图书情报领域

王晰巍,贾若男,王铎,等.图书情报领域人工智能的研究热点及发展趋势研究[J].图书情报工作,2019,63(1):70-80.

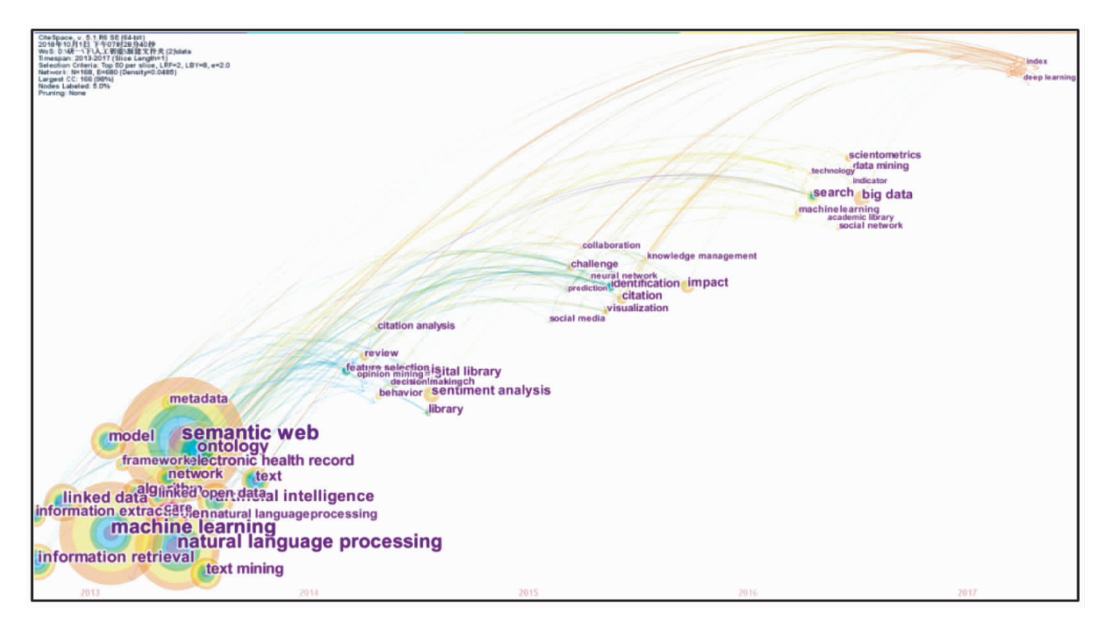


图6　国外图书情报领域研究关键词聚类知识图谱

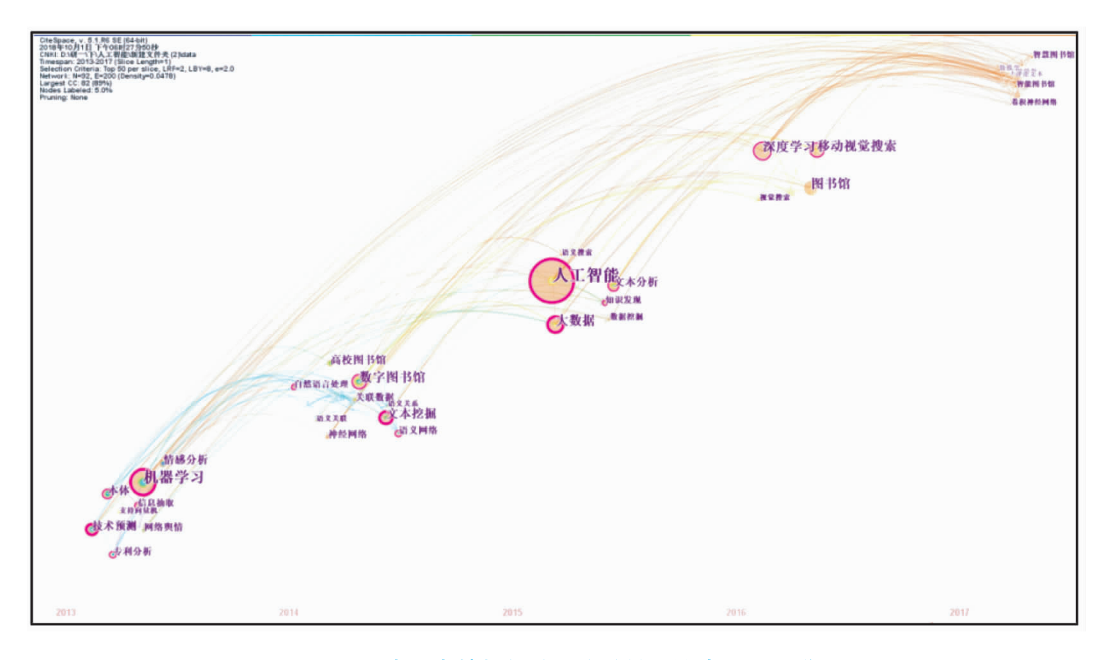


图7　国内图书情报领域研究关键词聚类知识图谱

的重要研究方向和趋势。人工智能技术在智慧图书馆建设中的应用研究在未来可以围绕两个方面进行:一方面,构建智慧图书馆智能管理体系。在人工和信息管理的基础上,将人工智能新技术与图书馆内容相结合,构建自动化文献检索定位系统、图书全帧系统、自动借阅系统等,对图书馆的文献资源、服务流程和管理模式进行智能化整合[50]。另一方面,提供智慧图书馆个性化智能推荐服务。根据读者的借阅、图书检索记录等信息,利用关联规则挖掘等技术或算法,找出读者的借阅规则和个性化信息需求,以个性化的方式为读者推荐图书或图书馆相关资讯服务信息[51]。另外,还可以从多个角度出发,利用人工智能的计算机视觉和智能语音等技术同图书馆资源连接起来,生成个性化智能推荐和语音应答服务系统,从而针对图书馆用户需求实现更为更精准、更为有效和更为智能的推荐服务[52]。

5.3　基于人工智能的社交网络数据挖掘与分析研究

随着智能时代的到来,信息、数据每分每秒都在以爆炸式的速度增长,加之社交网络的独特性,传统的数据处理技术已经不能满足数据的存储、处理和可视化分析等方面的要求。人工智能技术在社交网络数据挖掘与分析中将会发挥越来越重要的作用。具体的研究

可以从以下两个方面展开：一是对社交网络海量数据的可视化。利用自然语言处理技术和语义网对社交网络中的海量数据进行语义分析，并结合数据聚类和认知可视化技术来呈现数据信息的交互式可视化[53]；或者采用诸如蜂巢图等可视化方法，对社交网络中海量复杂数据的关系进行视觉化呈现，方便对社交网络的交互关系、信息传播路径等进行分析和发现[54]。二是对社交网络中的舆情传播进行预测和社群关系发现。基于复杂网络分析等方法，构建热点预测算法，对社交媒体数据进行分析并预测其传导的覆盖范围、影响力以及传播热点等[55]。

5.4　人工智能对隐私与安全的影响研究

随着语音识别、人脸识别、机器学习算法的日渐成熟，再加上网络空间、虚拟空间和交互场景的增加，使得个人数据更易于收集和分享，极大地方便了用户身份信息的编号，健康状态、信用记录、位置活动踪迹等信息的存储、分析和记录。与此同时，个人数据隐私泄露的途径与程度却很难通过网络进行追踪。因此，如何利用人工智能技术保障个人信息的隐私与使用安全，将会成为未来重要的研究趋势：一方面以人工智能为主要技术支撑的智慧医疗在使用过程中，对于病人的电子病历、医院使用私人数据的归属权限规范，需要进行相应的隐私保护技术[56]、规范和监管机制的研究；另一方面，人工智能的应用会使得"人机交互"关系发生改变，人机交互更加频繁，虚拟和真实随意切换，不可预测性与不可逆性会触发一系列潜在风险。人工智能技术可能被少数人用于欺诈等犯罪行为，基于不正当手段获取的"用户数据画像"，可能被别有用心之人冒充熟人进行欺骗，因而需要构建基于知识和模糊逻辑的模型来改善信息安全风险[57]。

在推进中国人工智能在图情领域的研究发展的过程中，还需要从政策、人文等角度积极参与人工智能全球治理机制的构建，在人工智能技术发展、风险防范、道德伦理、信息保护等方面发挥图情学科的特色。同时，积极推进人工智能在图情领域应用的产学研深度协同合作，推动学科之间的交叉融合，在原有的图情领域进行纵深化研究的基础上，加强运用计算机技术手段与金融、旅游、人文、医疗等相关学科的交叉及协同研究；并运用工程化思维将图书情报领域的人工智能研究与工程、制造领域相结合，从而不断丰富和延伸图情学科的研究方法，在纵深图情学科研究的基础上，在我国新时代管理和国家治理中发挥学科更大的作用，培养更多迎合社会发展的复合型和创新型人才。

参考文献：

[1] 石纯一. 人工智能原理[M]. 北京：清华大学出版社，1993.

[2] 麦肯锡. 2017 全球人工智能报告[EB/OL]. [2018-07-24]. http://www.360doc.com/content/17/0910/19/37263053_686042639.shtml.

[3] 猎豹全球智库. 2018 中国人工智能报告[EB/OL]. [2018-07-24]. https://www.jiemian.com/article/2005999.html.

[4] 清华大学中国科技政策研究中心. 中国人工智能报告 2018 [EB/OL]. [2018-07-24]. https://wemedia.ifeng.com/71852099/wemedia.shtml.

[5] STRONGMAN K T, BROWN R. Visual search with meaningful and non-meaningful material[J]. Quarterly journal of experimental psychology, 1966, 18(2):164-168.

[6] PRINZ W, ATAIAN D. Two components and two stages in search performance: a case study in visual search[J]. Acta psychologica, 1973, 37(4):255-277.

[7] SALTON G, YANG C S, YU C T. A theory of term importance in automatic text analysis[J]. Journal of the American Society for Information Science & Technology, 2014, 26(1):33-44.

[8] NARENDRA K S, PARATHASARANTHY K. Identification and control of dynamical systems using neural networks[J]. IEEE transactions on neural networks, 1990, 1(1):4-27.

[9] BNLSARI A. Some analytical solutions to the general approximation problem for feedforward neural networks[J]. Neural networks, 1993 (6): 991-996.

[10] MORRIS A. Expert systems for library and information services-A review[J]. Information processing & management, 1991, 27(6): 713-724.

[11] 张春博，丁堃，贾龙飞. 国际人工智能领域计量与可视化研究——基于 AAAI 年会论文的分析[J]. 图书情报工作，2012，56(22)：69-76.

[12] 王菲菲，赵旭. 论文层面计量视角下的核心文献价值测度与识别研究——以人工智能领域为例[J]. 情报理论与实践，2017，40(6)：91-98.

[13] SOCHER R, PERELYGIN A, WU J, et al. Recursive deep models for semantic compositionality over a sentiment treebank[C]//Proceedings of the 2013 conference on empirical methods in natural language processing. Stroudsburg: Association for Computational Linguistics, 2013:1631-1642.

[14] HINTON G, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups[J]. IEEE signal processing magazine, 2012, 29(6):82-97.

[15] KALCHBRENNER N, GREFENSTETTE E, LIUNSOM P. A convolutional neural network for modelling sentences[EB/OL]. [2018-07-24]. https://arxiv.org/pdf/404.2188.pdf.

[16] NESHATI M, FALLAHNEJAD Z, BEIGY H. On dynamicity of expert finding in community question answering[J]. Information processing & management, 2017, 53(5):1026-1042.

[17] LI G, LU T, DING X, et al. Predicting collaborative edits of questions and answers in online Q&A sites[J]. Journal of Internet technology, 2016, 17(6):1187－1194.

[18] 耿立大. 自然语言信息加工和人工智能[J]. 数据分析与知识发现,1983(1):30－35.

[19] 刘伟权,钟义信. 自然语言处理与全文情报检索[J]. 情报理论与实践,1997(1):43－46.

[20] 李蕾,郭祥昊. 基于语义网络的概念检索研究与实现[J]. 情报学报,2000,19(5):525－531.

[21] 刘贤锋. 基于 BP 神经网络预测企业竞争情报搜集成本[J]. 情报理论与实践,2007,30(5):646－649.

[22] 张秀华,赵伟. 基于径向基神经网络的数字馆藏质量评价研究[J]. 情报理论与实践,2009, 32(5):61－64.

[23] 张玉峰,晏创业. 基于机器学习的知识检索模型研究[J]. 图书情报知识,2002(4):6－9.

[24] 罗蔚. 数字化信息服务的个性化发展:从可适应到自适应[J]. 情报资料工作,2010,31(2):74－76.

[25] 逯万辉. 基于深度学习的学术期刊选题同质化测度方法研究[J]. 情报资料工作,2017(5): 105－112.

[26] 庞英智. Web 数据挖掘技术在电子商务中的应用[J]. 情报科学,2011(2):235－240.

[27] 邓仲华,刘伟伟, 陆颖隽. 基于云计算的大数据挖掘内涵及解决方案研究[J]. 情报理论与实践,2015,38(7):103－108.

[28] 陈臣. 基于大数据的图书馆个性化智慧服务体系构建[J]. 情报资料工作,2013(6):75－79.

[29] ZHANG R, DING G, ZHANG F, et al. The information retrieval technology of dynamic feedback artificial intelligence based on the neural network[C]// International conference on smart city and systems engineering. Washington:IEEE, 2017:250－253.

[30] ZHU Y J, YAN E J,SONG I Y. The use of a graph-based system to improve bibliographic information retrieval: system design, implementation, and evaluation[J]. Journal of The Association for Information Science and Technology,2017,68(2):480－490.

[31] RAHOMAN M M, ISHISE R. A proposal of a temporal semantics aware linked data information retrieval framework[J]. Journal of intelligent information systems, 2018, 50(3):573－595.

[32] ZOONEN W V, VAN DER MEER,TONI, GLA. Social media research: the application of supervised machine learning in organizational communication research[J]. Computers in human behavior, 2016, 63:132－141.

[33] LEE I. Social media analytics for enterprises: typology, methods, and processes[J]. Business horizons, 2018,61(2):119－210.

[34] YOO S Y, SONG J I, JEONG O R. Social media contents based sentiment analysis and prediction system[J]. Expert systems with applications, 2018,105:102－111.

[35] ATTIAOUI D, MARTIN A, YAGHLANE B B. Belief measure of expertise for experts detection in question answering communities: case study stack overflow[J]. Procedia computer science, 2017, 112:622－631.

[36] 侯治平,黄少杰,廉同辉,等. 基于语义关联分析的学术网络舆情传播研究——以科学网屠呦呦获诺贝尔奖博文为例[J]. 情报杂志, 2017, 36(5):118－123.

[37] 陈建英,朱晖,刘小芳. 基于灰色关联计算的网络民族舆情事件监测模型研究[J]. 情报杂志,2017,36(12):126－131,179.

[38] 吴鹏,刘恒旺,沈思. 基于深度学习和 OCC 情感规则的网络舆情情感识别研究[J]. 情报学报,2017,36(9):972－980.

[39] 李晨晖,张兴旺,秦晓珠. 基于大数据的文化遗产数字图书馆移动视觉搜索机制建设研究[J]. 情报理论与实践, 2018,41(4): 139－144,133.

[40] 齐云飞,赵宇翔,朱庆华. 关联数据在数字图书馆移动视觉搜索系统中的应用研究[J]. 数据分析与知识发现,2017,1(1):81－90.

[41] 张兴旺. 以信息推荐为例探讨图书馆人工智能体系的基本运作模式[J]. 情报理论与实践,2017,40(12):69－74.

[42] 李红巍. 基于本体相似度和时间衰减的动态个性化推荐算法[J]. 图书情报工作,2017,61(S1):95－98.

[43] 李慧,柴亚青. 基于属性特征的评论文本情感极性量化分析[J]. 数据分析与知识发现,2017,1(10):1－11.

[44] 齐林峰. 利用实体解析的跨社交媒体同一用户识别[J]. 图书情报工作,2017,61(6):107－114.

[45] 张继东. 移动社交网络环境下基于情景化偏好的用户行为感知研究[J]. 情报理论与实践,2017,40(1):110－114.

[46] 胡蓉,唐振贵,朱庆华. 混合需求驱动的文内视觉资源移动视觉搜索框架[J]. 情报学报,2018,37(3):285－293.

[47] GAO F,ZHANG X, HUANG Y, et al. Data-driven lightweight interest point selection for large scale visual search[J]. IEEE transactions on multimedia, 2018, 20(10): 2774－2787.

[48] MENEMENCIOGLU O,ORAK I M. An attempt at automatic label generation for object entry and exit on multimedia with a semantic search[J]. Tem journal-technology education management informatrics,2018,7(1):25－32.

[49] 夏立新,白阳,张心怡. 融合与重构:智慧图书馆发展新形态[J]. 中国图书馆学报,2018,44(1):35－49.

[50] GAO Y P. Implementation of an intelligent library system Based on WSN and RFID[J]. International journal of online engineering, 2018,14(5):221－224.

[51] YI K, CJEN T, CONG G. Library personalized recommendation service method based on improved association rules[J]. Library hi tech, 2018,36(3):443－457.

[52] SIMOVIC A. A big data smart library recommender system for an educational institution[J]. Library hi tech, 2018,36(3):498－523.

[53] WANG Y, BACIU G, LI C. Cognitive visualization of popular regions discovered from geo-tagged social media data[J]. International Journal of cognitive informatics & natural intelligence, 2018, 12(1):14－28.

[54] VERSPOOR K, OFOPHI B, GRANDA M R. CommViz: visualization of semantic patterns in large social communication networks

[J]. Information visualization, 2018,17(1):66-88.
[55] 王林森,王学义. 微博内向型传导热点发现与预测算法研究[J]. 图书情报工作,2018,62(3):71-77.
[56] WEI P, ZHOU Z. Research on security of information sharing in Internet of Things based on key algorithm[J]. Future generation computer systems, 2018,88:599-605.
[57] BASALLO Y A, SENTI V E, SANCHEZ N M. Artificial intelligence techniques for information security risk assessment[J]. IEEE Latin America transactions, 2018, 16(3):897-901.
[58] 胡刃锋. 产学研协同创新隐性知识共享机制[M]. 北京:光明日报出版社,2018.

作者贡献说明:

王晰巍:提出研究命题、研究思路及修订论文最后版本;
贾若男:负责论文撰写、修改及数据采集;
王铎:负责英文文献的收集及摘要的翻译;
郭宇:知识图谱可视化分析。

Research on the Development Trend of Artificial Intelligence Research in Library and Information Field

Wang Xiwei[1,2] Jia Ruonan[1] Wang Duo[1] Guo Yu[1]
[1] School of Management, Jilin University, Changchun 130022
[2] Research Center for Big Data Management, Jilin University, Changchun 130022

Abstract: [**Purpose/significance**] The article collects and analyzes relevant literatures on artificial intelligence research in the field of library and information at home and abroad, in order to understand the current research progress and future research trends of artificial intelligence in the field of library and information, and provide reference and research for academic and industrial research. [**Method/process**] Using literature analysis and knowledge map visualization method, the origin and development of artificial intelligence research in the field of library and information at home and abroad, domestic and foreign research hotspots and future research trends are compared and visualized. [**Result/conclusion**] From the perspective of research and development trends, the number of relevant research results at home and abroad is increasing year by year. From the research hotspots, foreign research mainly focuses on the application research of artificial intelligence in information retrieval, social media content classification and sentiment analysis, and knowledge question and answer. Domestic research mainly focuses on artificial intelligence in network public opinion analysis, library intelligent search and recommendation services, and social media network data analysis. From the research trend, future research should focus on the four main directions of artificial intelligence technology application visual search and semantic search, smart library service, social network data mining, and artificial intelligence impact on privacy and security.

Keywords: library and information artificial intelligence research hotspot development trend